Introduction to Epistemic Game Theory

Eric Pacuit

University of Maryland, College Park
`pacuit.org`
`epacuit@umd.edu`

May 25, 2016

Plan

- Introductory remarks
- Games and Game Models
- Backward and Forward Induction Reasoning
- Concluding remarks

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

dev.pacuit.org/games/avg

The Guessing Game, again

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

The Guessing Game, again

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

dev.pacuit.org/games/avg

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.
4. The person that wrote the smaller number will receive that amount plus 2 EUR (as a reward), and the person that wrote the larger number will receive the smaller number minus 2 EUR (as a punishment).

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.
4. The person that wrote the smaller number will receive that amount plus 2 EUR (as a reward), and the person that wrote the larger number will receive the smaller number minus 2 EUR (as a punishment).

Suppose that you are randomly paired with another person here at NASSLLI. What number would you write down?

dev.pacuit.org/games/td

From Decisions to Games, I

Commenting on the difference between Robin Crusoe's maximization problem and the maximization problem faced by participants in a social economy, von Neumann and Morgenstern write:

"Every participant can determine the variables which describe his own actions but not those of the others. **Nevertheless those "alien" variables cannot, from his point of view, be described by statistical assumptions**.

From Decisions to Games, I

Commenting on the difference between Robin Crusoe's maximization problem and the maximization problem faced by participants in a social economy, von Neumann and Morgenstern write:

"Every participant can determine the variables which describe his own actions but not those of the others. **Nevertheless those "alien" variables cannot, from his point of view, be described by statistical assumptions**. This is because the others are guided, just as he himself, by rational principles—whatever that may mean—and no *modus procedendi* can be correct which does not attempt to understand those principles and the interactions of the conflicting interests of all participants."

(vNM, pg. 11)

Just Enough Game Theory

A **game** is a mathematical model of a social interaction that includes

- the players (N);
- the actions (strategies) the players *can* take (for $i \in N$, $S_i \neq \emptyset$);
- the players' interests (for $i \in N$, $u_i : \Pi_{i \in N} S_i \rightarrow \mathbb{R}$); and
- the "structure" of the decision problem.

Just Enough Game Theory

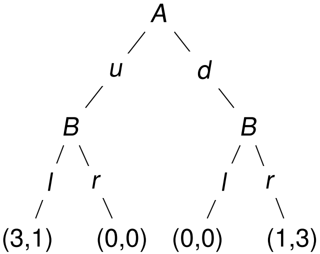
A **game** is a mathematical model of a social interaction that includes

- the players (N);
- the actions (strategies) the players *can* take (for $i \in N$, $S_i \neq \emptyset$);
- the players' interests (for $i \in N$, $u_i : \Pi_{i \in N} S_i \rightarrow \mathbb{R}$); and
- the "structure" of the decision problem.

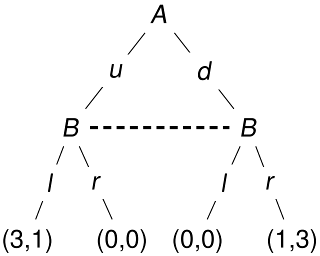
It does not specify the actions that the players do take.

		B	
		l	r
A	u	3, 1	0, 0
	d	0, 0	1, 3

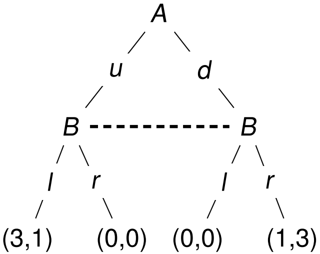
		B	
		l	r
A	u	3, 1	0, 0
	d	0, 0	1, 3

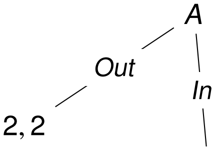


		B	
		l	r
A	u	3, 1	0, 0
	d	0, 0	1, 3



		B	
		l	r
A	u	3, 1	0, 0
	d	0, 0	1, 3





		B	
		l	r
A	u	3, 1	0, 0
	d	0, 0	1, 3

Just Enough Game Theory: Solution Concepts

A **solution concept** is a systematic description of the outcomes that may emerge in a family of games.

This is the starting point for most of game theory and includes many variants: Nash equilibrium (and refinements), backwards induction, or iterated dominance of various kinds.

They are usually thought of as the embodiment of “rational behavior” in some way and used to analyze game situations.

Game Models

- A game is a *partial* description of a set (or sequence) of interdependent **(Bayesian) decision problems**.

Game Models

- A game is a *partial* description of a set (or sequence) of interdependent **(Bayesian) decision problems**.

  A game will not normally contain enough information to determine what the players *believe* about each other.

Game Models

- A game is a *partial* description of a set (or sequence) of interdependent **(Bayesian) decision problems**.

  A game will not normally contain enough information to determine what the players *believe* about each other.

- A **model of a game** is a completion of the partial specification of the Bayesian decision problems *and* a representation of a particular play of the game.

Game Models

- A game is a *partial* description of a set (or sequence) of interdependent **(Bayesian) decision problems**.

  A game will not normally contain enough information to determine what the players *believe* about each other.

- A **model of a game** is a completion of the partial specification of the Bayesian decision problems *and* a representation of a particular play of the game.

- There are no special rules of rationality telling one what to do in the absence of degrees of belief except: decide what you believe, and then **maximize (subjective) expected utility**.

Knowledge and beliefs in game situations

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III*. *Management Science Theory* **14***: 159-182, 1967-68.*

R. Aumann. *Interactive Epistemology I & II*. International Journal of Game Theory (1999).

P. Battigalli and G. Bonanno. *Recent results on belief, knowledge and the epistemic foundations of game theory*. Research in Economics (1999).

R. Myerson. *Harsanyi's Games with Incomplete Information*. Special 50th anniversary issue of *Management Science*, 2004.

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

1. incomplete information: uncertainty about the *structure* of the game (outcomes, payoffs, strategy space)
2. imperfect information: uncertainty *within the game* about the previous moves of the players

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

Information in games situations

- Various states of information disclosure.

Information in games situations

- Various states of information disclosure.
    - *ex ante*, *ex interim*, *ex post*

Information in games situations

- Various states of information disclosure.
  - *ex ante*, *ex interim*, *ex post*

- Various “types” of information:

Information in games situations

- Various states of information disclosure.
  - *ex ante*, *ex interim*, *ex post*

- Various "types" of information:
  - imperfect information about the play of the game
  - incomplete information about the structure of the game
  - strategic information (what will the other players do?)
  - higher-order information (what are the other players thinking?)

Information in games situations

- Various states of information disclosure.
  - *ex ante*, *ex interim*, *ex post*

- Various “types” of information:
  - imperfect information about the play of the game
  - incomplete information about the structure of the game
  - strategic information (what will the other players do?)
  - higher-order information (what are the other players thinking?)

- Varieties of informational attitudes

Information in games situations

- Various states of information disclosure.
  - *ex ante*, *ex interim*, *ex post*

- Various “types” of information:
  - imperfect information about the play of the game
  - incomplete information about the structure of the game
  - strategic information (what will the other players do?)
  - higher-order information (what are the other players thinking?)

- Varieties of informational attitudes
  - hard (“knowledge”)
  - soft (“beliefs”)

Epistemic Game Theory

Formally, a game is described by its strategy sets and payoff functions.

Epistemic Game Theory

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.

Epistemic Game Theory

> Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game....
>
> The difference lies in the attitudes of the players, in their expectations about each other, in custom, and in history, though the rules of the game do not distinguish between the two situations. (pg. 72)

R. Aumann and J. H. Dreze. *Rational Expectations in Games*. American Economic Review 98 (2008), pp. 72-86.

The Epistemic Program in Game Theory

...the analysis constitutes a fleshing-out of the textbook interpretation of equilibrium as ‘rationality plus correct beliefs.’

E. Dekel and M. Siniscalchi. *Epistemic Game Theory*. Handbook of Game Theory with Economic Applications, 2015.

A. Brandenburger. *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*. World Scientific, 2014.

A. Perea. *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press, 2012.

EP and O. Roy. *Epistemic Game Theory*. Stanford Encyclopedia of Philosophy, 2015.

Bayesian Decision Problem

	it rains	it does not rain
take umbrella	encumbered, dry	encumbered, dry
leave umbrella	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

Expected utility of action *A*

$$EU(A) = \sum_{o\in O} P_A(o) \times U(o)$$

Expected utility of action *A*

Utility of outcome *o*

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

Expected utility of action *A*

Utility of outcome *o*

Probability of outcome *o* conditional on *A*

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \;\&\; A)}{P(A)}$

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \,\&\, A)}{P(A)}$

Classical: $P_A(o) = \sum_{s \in S} P(s) f_{A,s}(o)$, where

$$f_{A,s}(o) = \begin{cases} 1 & A(s) = o \\ 0 & A(s) \neq o \end{cases}$$

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \,\&\, A)}{P(A)}$

Classical: $P_A(o) = \sum_{s \in S} P(s) f_{A,s}(o)$, where

$$f_{A,s}(o) = \begin{cases} 1 & A(s) = o \\ 0 & A(s) \neq o \end{cases}$$

Causal: $P_A(o) = P(A \mathrel{\Box\!\!\rightarrow} o)$

P(“if A were performed, outcome o would ensue”)

(Lewis, 1981)

		B	
		L	R
A	U	2,1	0,0
	D	0,0	1,2

		$\text{Play}_B(L)$	$\text{Play}_B(R)$
A	U	2	0
	D	0	1

		$\text{Play}_A(U)$	$\text{Play}_A(D)$
B	L	1	0
	R	0	2

Models of Games

Suppose that G is a game.

- Outcomes of the game: $S = \Pi_{i \in N} S_i$
- Player i's partial beliefs (or conjecture): $P_i \in \Delta(S_{-i})$

  $\Delta(X)$ *is the set of probabilities measures over* X

Models of Games, continued

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$ is a strategic (form of a) game.

Models of Games, continued

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$ is a strategic (form of a) game.

- W is a set of *possible worlds* (possible outcomes of the game)
- **s** is a function $\mathbf{s} : W \rightarrow \Pi_{i \in N} S_i$

Models of Games, continued

$G = \langle N, \{S_i, u_i\}_{i\in N}\rangle$ is a strategic (form of a) game.

- W is a set of *possible worlds* (possible outcomes of the game)
- **s** is a function $\mathbf{s} : W \to \Pi_{i\in N} S_i$
- For $s_i \in S_i$, $[s_i] = \{w \mid \mathbf{s}(w) = s_i\}$, if $X \subseteq S$, $[X] = \bigcup_{s\in X}[s]$.

Models of Games, continued

$G = \langle N, \{S_i, u_i\}_{i \in N} \rangle$ is a strategic (form of a) game.

- W is a set of *possible worlds* (possible outcomes of the game)
- $\mathbf{s}$ is a function $\mathbf{s} : W \rightarrow \Pi_{i \in N} S_i$
- For $s_i \in S_i$, $[s_i] = \{w \mid \mathbf{s}(w) = s_i\}$, if $X \subseteq S$, $[X] = \bigcup_{s \in X}[s]$.
- ***ex ante* beliefs**: For each $i \in N$, $P_i \in \Delta(W)$. Two assumptions:
  - $[s]$ is measurable for all strategy profiles $s \in S$
  - $P_i([s_i]) > 0$ for all $s_i \in S_i$

ex interim **beliefs**: $P_{i,w} \in \Delta(S_{-i})$

- ...given player *i*'s choice: $P_{i,w}(\cdot) = P_i(\cdot \mid [\mathbf{s}_i(w)])$
- ...given all player *i knows*: $P_{i,w}(\cdot) = P_i(\cdot \mid K_i)$, $\qquad K_i \subseteq [\mathbf{s}_i(w)]$
- ...given all player *i fully believes*: $P_{i,w}(\cdot) = P_i(\cdot \mid B_i)$, $\quad B_i \subseteq [\mathbf{s}_i(w)]$

ex interim **beliefs**: $P_{i,w} \in \Delta(S_{-i})$

- ...given player *i*'s choice: $P_{i,w}(\cdot) = P_i(\cdot \mid [\mathbf{s}_i(w)])$
- ...given all player *i knows*: $P_{i,w}(\cdot) = P_i(\cdot \mid K_i)$, $\qquad K_i \subseteq [\mathbf{s}_i(w)]$
- ...given all player *i fully believes*: $P_{i,w}(\cdot) = P_i(\cdot \mid B_i)$, $\quad B_i \subseteq [\mathbf{s}_i(w)]$

Expected utility: Given $P \in \Delta(S_{-i})$,

$$EU_{i,P}(a) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(a, s_{-i})$$

ex interim **beliefs**: $P_{i,w} \in \Delta(S_{-i})$

- ...given player *i*'s choice: $P_{i,w}(\cdot) = P_i(\cdot \mid [\mathbf{s}_i(w)])$
- ...given all player *i knows*: $P_{i,w}(\cdot) = P_i(\cdot \mid K_i), \qquad K_i \subseteq [\mathbf{s}_i(w)]$
- ...given all player *i fully believes*: $P_{i,w}(\cdot) = P_i(\cdot \mid B_i), \quad B_i \subseteq [\mathbf{s}_i(w)]$

Expected utility: Given $w \in W$,

$$EU_{i,w}(a) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}])u_i(a, s_{-i})$$

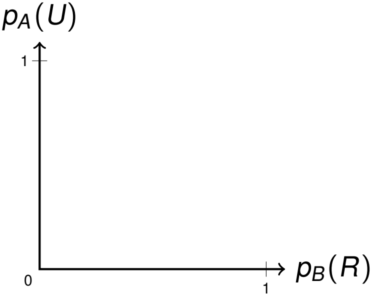
B

		L	*R*
A	*U*	2,1	0,0
	D	0,0	1,2

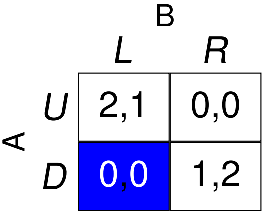
Strategic Game

		B	
		L	R
A	U	2,1	0,0
	D	0,0	1,2

Strategic Game

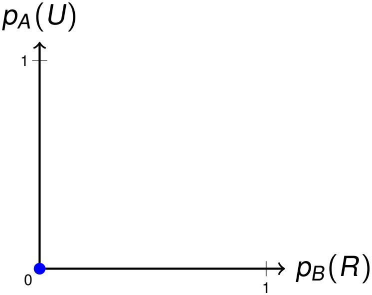


Solution Space



		B	
		L	R
A	U	2,1	0,0
	D	0,0	1,2

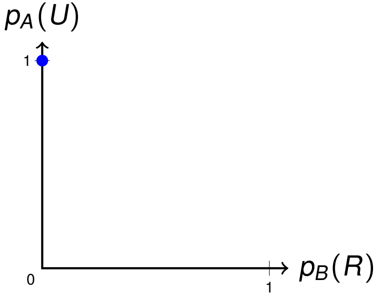
Strategic Game



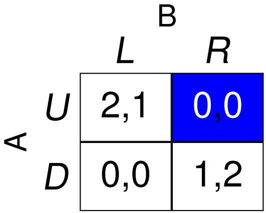
Solution Space

		B	
		L	R
A	U	2,1	0,0
	D	0,0	1,2

Strategic Game



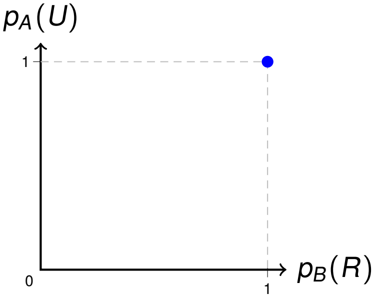
Solution Space



B

		L	R
A	U	2,1	0,0
	D	0,0	1,2

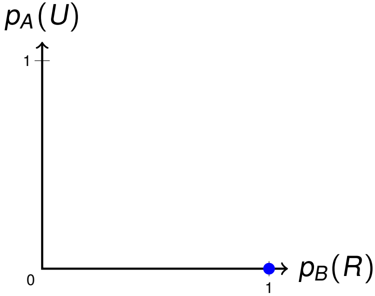
Strategic Game



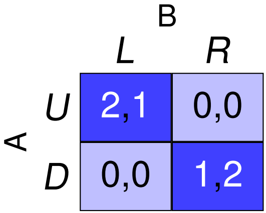
Solution Space

		B	
		L	*R*
A	*U*	2,1	0,0
	D	0,0	1,2

Strategic Game

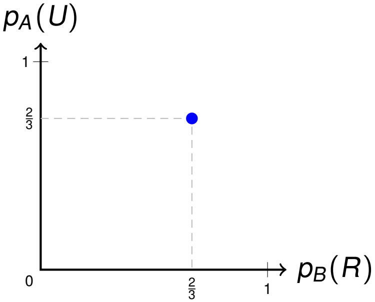


Solution Space



		B	
		L	R
A	U	2,1	0,0
	D	0,0	1,2

Strategic Game



Solution Space

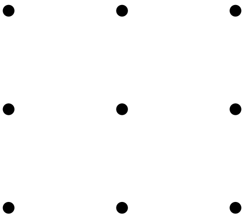
An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1

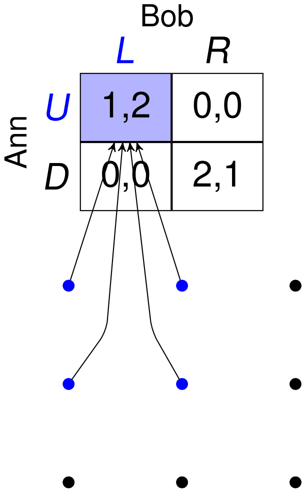
An Example

Bob

Ann	*L*	*R*
U	1,2	0,0
D	0,0	2,1



An Example

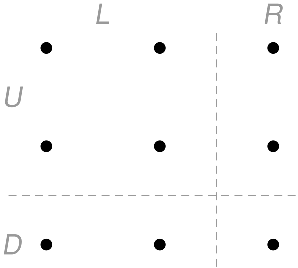


		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1

An Example

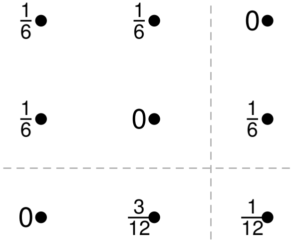
Bob

		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



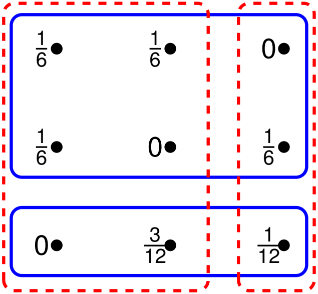
An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



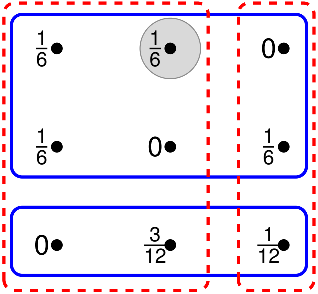
An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



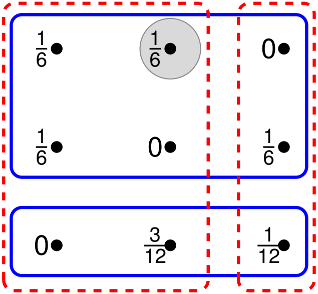
An Example

		Bob	
		L	R
Ann	U	1,2	0,0
	D	0,0	2,1



An Example

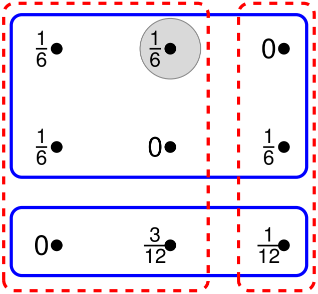
		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



- Ann's choice is *optimal* (given her information)

An Example

		Bob	
		L	R
Ann	U	1,2	0,0
	D	0,0	2,1

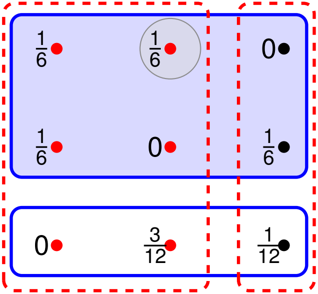


- Ann's choice is *optimal* (given her information)

$$1 \cdot P_A(L) + 0 \cdot P_A(R) \geq 0 \cdot P_A(L) + 2 \cdot P_A(R)$$

An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1

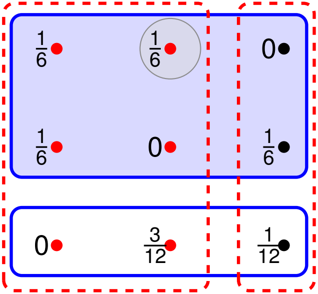


- Ann's choice is *optimal* (given her information)

$$1 \cdot P_A(L) + 0 \cdot P_A(R) \geq 0 \cdot P_A(L) + 2 \cdot P_A(R)$$

An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1

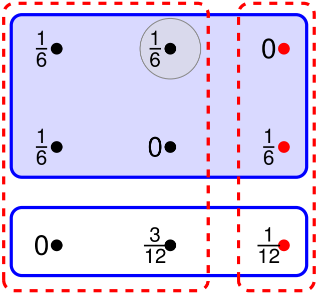


- Ann's choice is *optimal* (given her information)

$$1 \cdot \frac{3}{4} + 0 \cdot P_A(R) \geq 0 \cdot \frac{3}{4} + 2 \cdot P_A(R)$$

An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



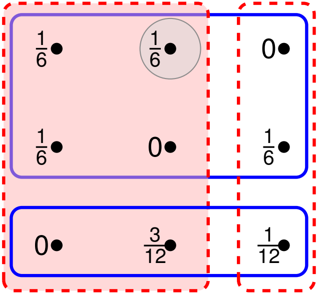
- Ann's choice is *optimal* (given her information)

$$1 \cdot \frac{3}{4} + 0 \cdot \frac{1}{4} \geq 0 \cdot \frac{3}{4} + 2 \cdot \frac{1}{4}$$

An Example

		Bob	
		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1

- Ann's choice is *optimal* (given her information)
- Bob's choice is *optimal* (given his information)

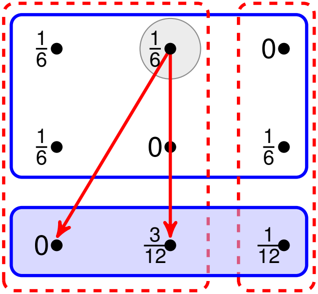


$$2 \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} \geq 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}$$

An Example

Bob

		L	*R*
Ann	*U*	1,2	0,0
	D	0,0	2,1



- Ann's choice is *optimal* (given her information)
- Bob's choice is *optimal* (given his information)
- Bob *considers it possible* Ann is *irrational*

$$0 \cdot \frac{3}{4} + 2 \cdot \frac{1}{4} \not\geq 1 \cdot \frac{3}{4} + 0 \cdot \frac{1}{4}$$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_i, s_{-i})$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_i, s_{-i})$

$\text{Rat}_i = \{w \mid EU_{i,w}(\mathbf{s}_i(w)) \geq EU_{i,w}(s_i) \text{ for all } s_i \in S_i\}$

For any $P \in \Delta(S_{-i})$ and $s_i \in S_i$, $EU_{i,P}(s_i) = \sum_{s_{-i} \in S_{-i}} P(s_{-i}) u_i(s_i, s_{-i})$

For any $w \in W$ and $s_i \in S_i$, $EU_{i,w}(s_i) = \sum_{s_{-i} \in S_{-i}} P_{i,w}([s_{-i}]) u_i(s_i, s_{-i})$

$\mathrm{Rat}_i = \{w \mid EU_{i,w}(\mathbf{s}_i(w)) \geq EU_{i,w}(s_i) \text{ for all } s_i \in S_i\}$

Each $P \in \Delta(W)$ is associated with $P^S \in \Delta(S)$ as follows: for all $s \in S$, $P^S(s) = P([s])$

Mixed Strategy

A **mixed strategy** is a sequence of probabilities over the players' available actions: $\sigma \in \Pi_{i \in N} \Delta(S_i)$, $P_\sigma \in \Delta(S)$.

Choices are assumed to be independent: $P_\sigma(s) = \sigma_1(s_1) \cdots \sigma_n(s_n)$

Mixed Strategies

“We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason for each action we take. Outside of Las Vegas we do not spin roulettes.”

Mixed Strategies

“We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason for each action we take. Outside of Las Vegas we do not spin roulettes.”

- One can think about a game as an interaction between large populations...a mixed strategy is viewed as the distribution of the pure choices in the population.

Mixed Strategies

“We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason for each action we take. Outside of Las Vegas we do not spin roulettes.”

- One can think about a game as an interaction between large populations...a mixed strategy is viewed as the distribution of the pure choices in the population.
- *Harsanyi’s purification theorem*: A player’s mixed strategy is thought of as a plan of action which is dependent on private information which is not specified in the model. Although the player’s behavior appears to be random, it is actually deterministic.

Mixed Strategies

"We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason for each action we take. Outside of Las Vegas we do not spin roulettes."

- One can think about a game as an interaction between large populations...a mixed strategy is viewed as the distribution of the pure choices in the population.
- *Harsanyi's purification theorem*: A player's mixed strategy is thought of as a plan of action which is dependent on private information which is not specified in the model. Although the player's behavior appears to be random, it is actually deterministic.
- Mixed strategies are beliefs held by all *other* players concerning a player's actions.

Nash Equilibrium

Let $\langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game

For $s_{-i} \in S_{-i}$, let

$$B_i(s_{-i}) = \{s_i \in S_i \mid u_i(s_{-i}, s_i) \geq u_i(s_{-i}, s_i') \ \forall \ s_i' \in S_i\}$$

B_i is the **best-response** function.

Nash Equilibrium

Let $\langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ be a strategic game

For $s_{-i} \in S_{-i}$, let

$$B_i(s_{-i}) = \{s_i \in S_i \mid u_i(s_{-i}, s_i) \geq u_i(s_{-i}, s_i') \ \forall\ s_i' \in S_i\}$$

B_i is the **best-response** function.

$s^* \in S$ is a **Nash equilibrium** iff $s_i^* \in B_i(s_{-i}^*)$ for all $i \in N$.

Nash Equilibrium: Example

	L	*R*
U	2,1	0,0
D	0,0	1,2

Nash Equilibrium: Example

	L	R
U	2,1	0,0
D	0,0	1,2

$$N = \{r, c\} \quad S_r = \{U, D\},\ S_c = \{L, R\}$$

Nash Equilibrium: Example

	L	R
U	2,1	0,0
D	0,0	1,2

$$N = \{r, c\} \quad S_r = \{U, D\},\ S_c = \{L, R\}$$

$$B_r(L) = \{U\} \qquad\qquad B_r(R) = \{D\}$$

Nash Equilibrium: Example

	L	R
U	2,1	0,0
D	0,0	1,2

$$N = \{r, c\} \quad S_r = \{U, D\},\ S_c = \{L, R\}$$

$$B_r(L) = \{U\} \qquad B_r(R) = \{D\}$$

$$B_c(U) = \{L\} \qquad B_c(D) = \{R\}$$

Nash Equilibrium: Example

	L	R
U	2,1	0,0
D	0,0	1,2

$$N = \{r, c\} \quad S_r = \{U, D\},\ S_c = \{L, R\}$$

$$B_r(L) = \{U\} \qquad\qquad B_r(R) = \{D\}$$

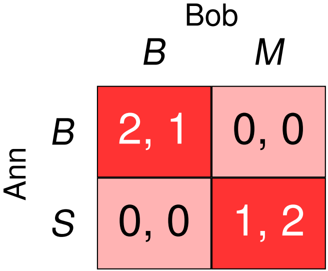
$$B_c(U) = \{L\} \qquad\qquad B_c(D) = \{R\}$$

(U, L) is a Nash Equilibrium $\qquad$ (D, R) is a Nash Equilibrium

Battle of the Sexes

		Bob	
		B	*S*
Ann	*B*	2, 1	0, 0
	S	0, 0	1, 2

Battle of the Sexes



		Bob	
		B	M
Ann	B	2, 1	0, 0
	S	0, 0	1, 2

(B, B) (S, S), and $([2/3 : B, 1/3 : S], [1/3 : B, 2/3 : S])$ are Nash equilibria.

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess?

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? 100

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, 99

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , 67

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , ~~67~~, . . . , 2, 1

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , ~~67~~, . . . , ~~2~~, ①

Traveler's Dilemma

$$\begin{array}{c|cccccc}
 & 2 & 3 & 4 & \cdots & 99 & 100 \\
\hline
2 & (2,2) & (4,0) & (4,0) & \cdots & (4,0) & (4,0) \\
3 & (0,4) & (3,3) & (5,1) & \cdots & (5,1) & (5,1) \\
4 & (0,4) & (1,5) & (4,4) & \cdots & (6,2) & (6,2) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
99 & (0,4) & (1,5) & (2,6) & \cdots & (99,99) & (101,97) \\
100 & (0,4) & (1,5) & (2,6) & \cdots & (97,101) & (100,100)
\end{array}$$

Traveler's Dilemma

	2	3	4	⋯	99	100
2	$(2, 2)$	$(4, 0)$	$(4, 0)$	⋯	$(4, 0)$	$(4, 0)$
3	$(0, 4)$	$(3, 3)$	$(5, 1)$	⋯	$(5, 1)$	$(5, 1)$
4	$(0, 4)$	$(1, 5)$	$(4, 4)$	⋯	$(6, 2)$	$(6, 2)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	$(0, 4)$	$(1, 5)$	$(2, 6)$	⋯	$(99, 99)$	$(101, 97)$
100	$(0, 4)$	$(1, 5)$	$(2, 6)$	⋯	$(97, 101)$	$(100, 100)$

$(2, 2)$ is the only Nash equilibrium.

Traveler’s Dilemma

	2	3	4	⋯	99	100
2	(2, 2)	(4, 0)	(4, 0)	⋯	(4, 0)	(4, 0)
3	(0, 4)	(3, 3)	(5, 1)	⋯	(5, 1)	(5, 1)
4	(0, 4)	(1, 5)	(4, 4)	⋯	(6, 2)	(6, 2)
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	(0, 4)	(1, 5)	(2, 6)	⋯	(99, 99)	(101, 97)
100	(0, 4)	(1, 5)	(2, 6)	⋯	(97, 101)	(100, 100)

The analysis is insensitive to the amount of reward/punishment.

Characterizing Nash Equilibria

Theorem (Aumann). σ is a Nash equilibrium of G iff there exists a model $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$ such that:

- for all $i \in N$, $\text{Rat}_i = W$;
- for all $i, j \in N$, $P_i = P_j$; and
- for all $i \in N$, $P_i^S = P_\sigma$.

Playing a Nash equilibrium is *required* by the players rationality and *common knowledge* thereof.

- Players need not be *certain* of the other players' beliefs
- Strategies that are not an equilibrium may be *rationalizable*
- Sometimes considerations of riskiness trump the Nash equilibrium

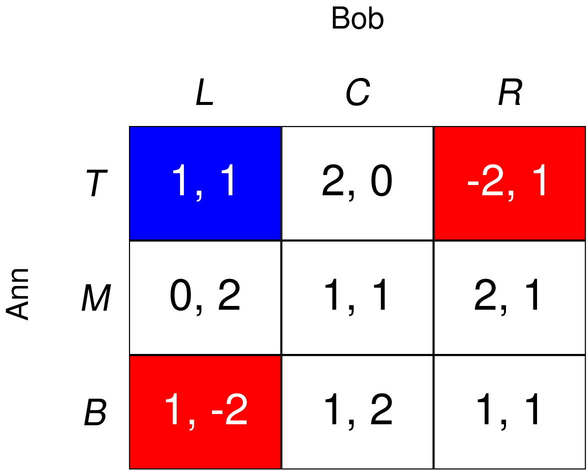
Bob

		L	C	R
Ann	T	1, 1	2, 0	-2, 1
	M	0, 2	1, 1	2, 1
	B	1, -2	1, 2	1, 1

Bob

Ann	*L*	*C*	*R*
T	1, 1	2, 0	-2, 1
M	0, 2	1, 1	2, 1
B	1, -2	1, 2	1, 1

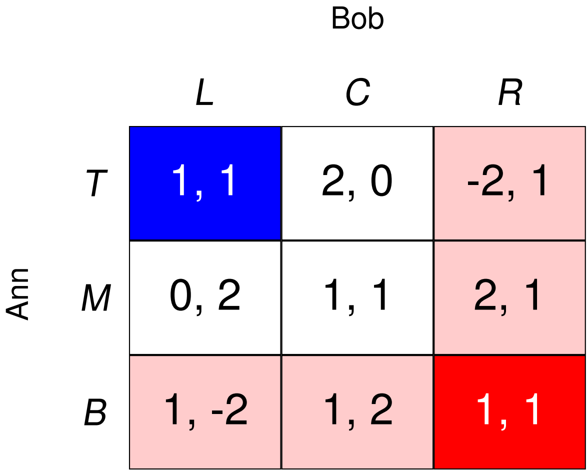
(T, L) is the unique pure-strategy Nash equilibrium



Bob

		L	*C*	*R*
Ann	*T*	1, 1	2, 0	-2, 1
	M	0, 2	1, 1	2, 1
	B	1, -2	1, 2	1, 1

(T, L) is the unique pure-strategy Nash equilibrium



Bob

Ann

	L	*C*	*R*
T	1, 1	2, 0	-2, 1
M	0, 2	1, 1	2, 1
B	1, -2	1, 2	1, 1

Why not play *B* and *R*?

Rationalizability

Bob

Ann

	L	*C*	*R*
T	3, 2	0, 0	2, 3
M	0, 0	1, 1	0, 0
B	2, 3	0, 0	3, 2

Bob

		L	C	R
Ann	T	3, 2	0, 0	2, 3
	M	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

(M, C) is the unique Nash equilibrium

		Bob		
		L	*C*	*R*
Ann	*T*	3, 2	0, 0	2, 3
	M	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

T, *L*, *B* and *R* are **rationalizable**

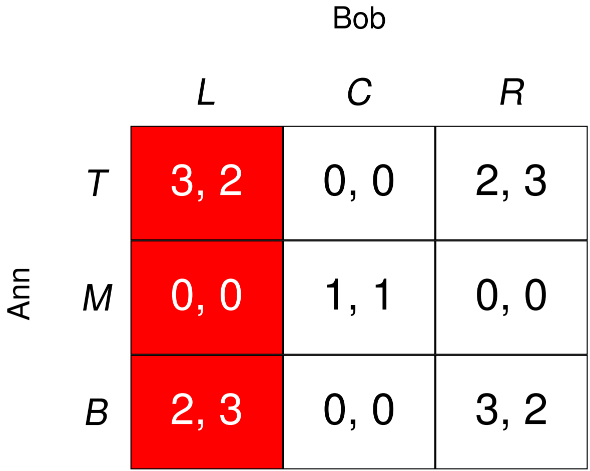
Bob

		L	*C*	*R*
	T	3, 2	0, 0	2, 3
Ann	*M*	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

T, *L*, *B* and *R* are **rationalizable**

		Bob		
		L	*C*	*R*
Ann	*T*	3, 2	0, 0	2, 3
	M	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

Ann plays *B* because she thought Bob will play *R*



		Bob		
		L	*C*	*R*
Ann	*T*	3, 2	0, 0	2, 3
	M	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

Bob plays *L* because she thought Ann will play *B*

		Bob		
		L	*C*	*R*
Ann	*T*	3, 2	0, 0	2, 3
	M	0, 0	1, 1	0, 0
	B	2, 3	0, 0	3, 2

Bob was correct, but Ann was wrong

		Bob			
		L	*C*	*R*	*X*
Ann	*T*	3, 2	0, 0	2, 3	0, -5
	M	0, 0	1, 1	0, 0	200,-5
	B	2, 3	0, 0	3, 2	1,-3

Not every strategy is rationalizable: Ann can't play *M because* she thinks Bob will play *X*

"Analysis of strategic economic situations requires us, implicitly or explicitly, to maintain as plausible certain psychological hypotheses. The hypothesis that real economic agents universally recognize the salience of Nash equilibria may well be less accurate than, for example, the hypothesis that agents attempt to "out-smart" or "second-guess" each other, believing that their opponents do likewise." (pg. 1010)

B. D. Bernheim. *Rationalizable Strategic Behavior*. Econometrica, 52:4, pgs. 1007 - 1028, 1984.

“The rules of a game and its numerical data are seldom sufficient for logical deduction alone to single out a unique choice of strategy for each player. *To do so one requires either richer information (such as institutional detail or perhaps historical precedent for a certain type of behavior) or bolder assumptions about how players choose strategies.* Putting further restrictions on strategic choice is a complex and treacherous task. But one’s intuition frequently points to patterns of behavior that cannot be isolated on the grounds of consistency alone.”

(pg. 1035)

D. G. Pearce. *Rationalizable Strategic Behavior*. Econometrica, 52, 4, pgs. 1029 - 1050, 1984.

Rationalizability

A **best reply set** (BRS) is a sequence $(B_1, B_2, \ldots, B_n) \subseteq S = \Pi_{i \in N} S_i$ such that for all $i \in N$, and all $s_i \in B_i$, there exists $\mu_{-i} \in \Delta(B_{-i})$ such that s_i is a best response to μ_{-i}: I.e.,

$$b_i = \arg\max_{s_i \in S_i} EU_{i,\mu_{-i}}(s_i)$$

.

2

		b_1	b_2	b_3	b_4
	a_1	0, 7	2, 5	7, 0	0, 1
1	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

		2			
		b_1	b_2	b_3	b_4
1	a_1	0, 7	2, 5	7, 0	0, 1
	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS

		2			
		b_1	b_2	b_3	b_4
	a_1	0, 7	2, 5	7, 0	0, 1
1	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS
- $(\{a_1, a_3\}, \{b_1, b_3\})$ is a BRS

2

		b_1	b_2	b_3	b_4
	a_1	0, 7	2, 5	7, 0	0, 1
1	a_2	5, 2	3, 3	5, 2	0, 1
	a_3	7, 0	2, 5	0, 7	0, 1
	a_4	0, 0	0, -2	0, 0	10, -1

- (a_2, b_2) is the unique Nash equilibria, hence $(\{a_2\}, \{b_2\})$ is a BRS
- $(\{a_1, a_3\}, \{b_1, b_3\})$ is a BRS
- $(\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\})$ is a full BRS

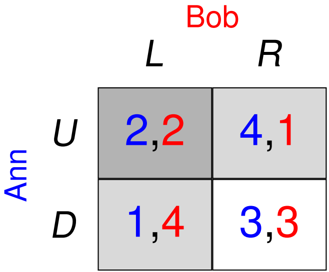
Theorem (Bernheim; Pearce; Brandenburger and Dekel; . . .).
$(B_1, B_2, \ldots, B_n)$ is a BRS for G iff there exists a model such that the *projection* of the the event where there is common belief that all players are rational is $B_1 \times \cdots \times B_n$.

		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2



		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* strictly dominates *D* and *L* strictly dominates *R*.

		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* strictly dominates *D* and *L* strictly dominates *R*.

Game 2: *U* strictly dominates *D*

		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* strictly dominates *D* and *L* strictly dominates *R*.

Game 2: *U* strictly dominates *D*, and *after removing D*, *L* strictly dominates *R*.

		Bob	
		L	*R*
Ann	*U*	2,2	4,1
	D	1,4	3,3

Game 1

		Bob	
		L	*R*
Ann	*U*	2,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* strictly dominates *D* and *L* strictly dominates *R*.

Game 2: *U* strictly dominates *D*, and *after removing D*, *L* strictly dominates *R*.

Theorem. In all models where the players are *rational* and there is *common belief of rationality*, the players choose strategies that survive iterative removal of strictly dominated strategies (and, conversely...).

		Bob	
		L	*R*
Ann	*U*	3,3	1,1
	D	2,2	2,2

Is *R* rationalizable?

		Bob	
		L	*R*
Ann	*U*	3,3	1,1
	D	2,2	2,2

Is *R* rationalizable?

There is no *full support* probability such that *R* is a best response

		Bob	
		L	*R*
Ann	*U*	3,3	1,1
	D	2,2	2,2

Is *R* rationalizable?

There is no *full support* probability such that *R* is a best response

Should Ann assign probability 0 to *R* or probability $>$ 0 to *R*?

Strategic Reasoning and Admissibility

"The argument for deletion of a weakly dominated strategy for player *i* is that he contemplates the possibility that every strategy combination of his rivals occurs with positive probability. However, this hypothesis clashes with the logic of iterated deletion, which assumes, precisely, that eliminated strategies are not expected to occur."

Mas-Colell, Whinston and Green. *Introduction to Microeconomics*. 1995.

A Puzzle

R. Cubitt and R. Sugden. *Rationally Justifiable Play and the Theory of Non-cooperative games*. Economic Journal, 104, pgs. 798 - 803, 1994.

R. Cubitt and R. Sugden. *Common reasoning in games: A Lewisian analysis of common knowledge of rationality*. Manuscript, 2011.

		Bob	
		L	*R*
Ann	*U*	1,1	0,0
	D	0,0	0,0

Game 1

		Bob	
		L	*R*
Ann	*U*	1,1	1,0
	D	1,0	0,1

Game 2

		Bob	
		L	*R*
Ann	*U*	1,1	0,0
	D	0,0	0,0

Game 1

		Bob	
		L	*R*
Ann	*U*	1,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

		Bob	
		L	*R*
Ann	*U*	1,1	0,0
	D	0,0	0,0

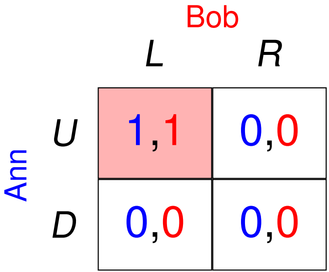
Game 1

		Bob	
		L	*R*
Ann	*U*	1,1	1,0
	D	1,0	0,1

Game 2

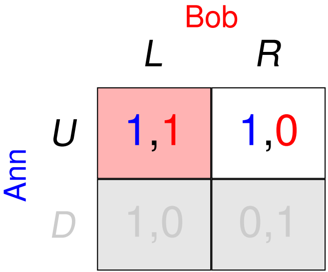
Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: *U* weakly dominates *D*



	L	*R*
U	1,1	0,0
D	0,0	0,0

Game 1



	L	*R*
U	1,1	1,0
D	1,0	0,1

Game 2

Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: *U* weakly dominates *D*, and *after removing D*, *L* strictly dominates *R*.

		Bob	
		L	*R*
Ann	*U*	1,1	0,0
	D	0,0	0,0

Game 1

		Bob	
		L	*R*
Ann	*U*	1,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: But, now what is the reason for not playing *D*?

Bob

		L	*R*
Ann	*U*	1,1	0,0
	D	0,0	0,0

Game 1

Bob

		L	*R*
Ann	*U*	1,1	1,0
	D	1,0	0,1

Game 2

Game 1: *U* weakly dominates *D* and *L* weakly dominates *R*.

Game 2: But, now what is the reason for not playing *D*?

Theorem (Samuelson). There is no model of Game 2 satisfying common knowledge of rationality (where rationality incorporates weak dominance).

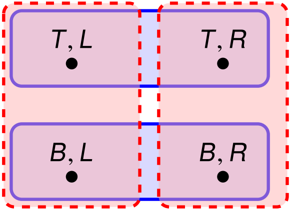
Common Knowledge of Admissibility

		Bob	
		L	*R*
Ann	*T*	1,1	1,0
	B	1,0	0,1

There is no model of this game with *common knowledge* of admissibility.

Common Knowledge of Admissibility

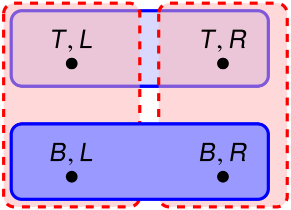
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



The ”full” model of the game

Common Knowledge of Admissibility

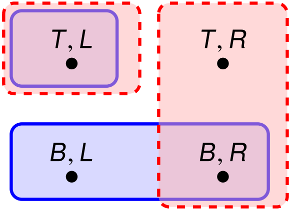
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



The ”full” model of the game: *B is not admissible given Ann’s information*

Common Knowledge of Admissibility

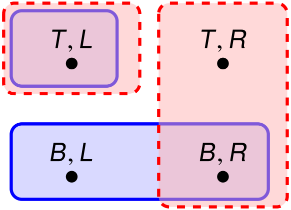
		Bob	
		L	R
Ann	T	1,1	1,0
	B	1,0	0,1



What is wrong with this model?

Common Knowledge of Admissibility

		Bob	
		L	*R*
Ann	*T*	1,1	1,0
	B	1,0	0,1



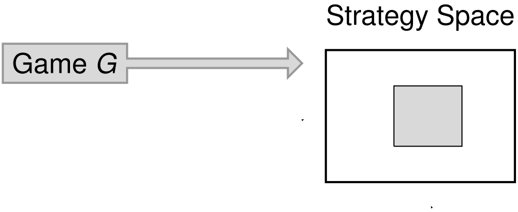
Privacy of Tie-Breaking/No Extraneous Beliefs: If a strategy is *rational* for an opponent, then it cannot be “ruled out”.

The Epistemic Program in Game Theory

Game *G*

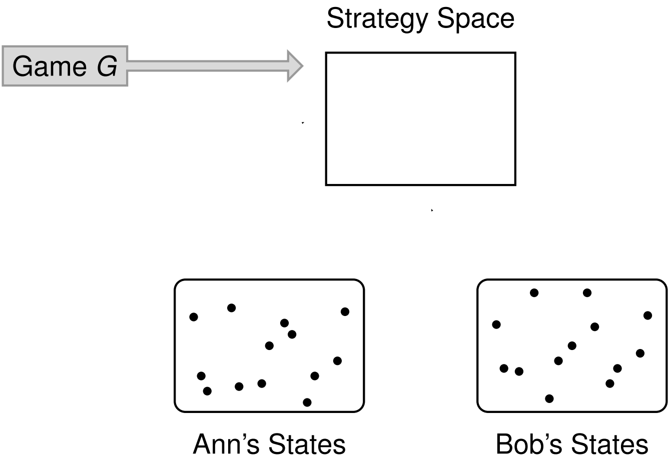
G: available actions, payoffs, structure of the decision problem

The Epistemic Program in Game Theory



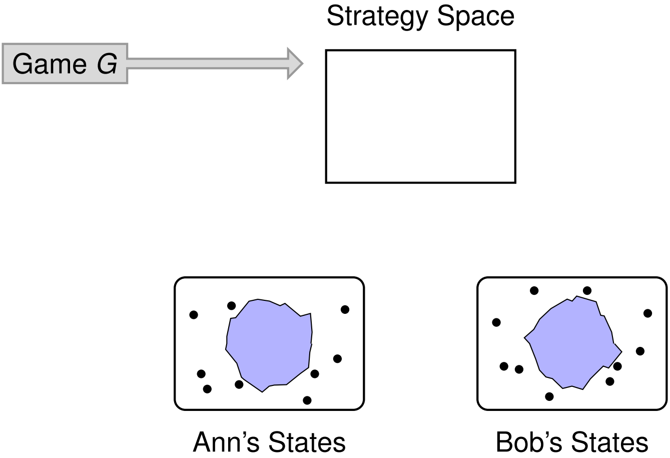
solution concepts are systematic descriptions of what players *do*

The Epistemic Program in Game Theory



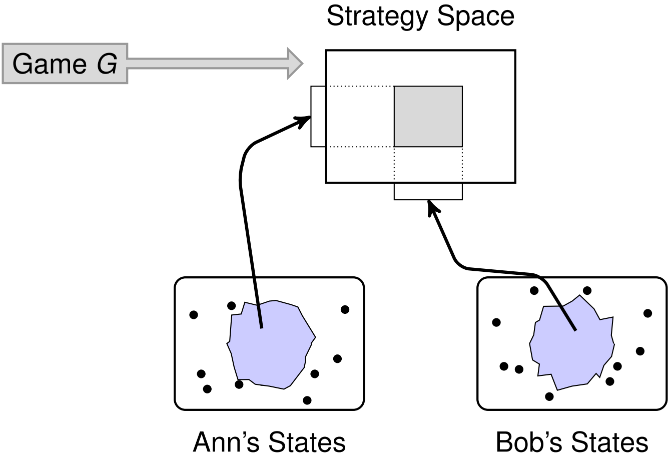
The game model includes *information states* of the players

The Epistemic Program in Game Theory



Restrict to information states satisfying some rationality condition

The Epistemic Program in Game Theory



Strategy Space

Game *G*

Ann's States

Bob's States

Project onto the strategy space

Why go beyond the basic Bayesian model?

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.

  "Even if I think you *know* what I am going to do, I can consider how I think you would react if I did something that you and I both know I will not do, and my answers to such counterfactual questions will be relevant to assessing the rationality of what I *am* going to do." (pg. 31)

  R. Stalnaker. *Belief revision in games: forward and backward induction*. Mathematical Social Sciences, 36, pgs. 31 - 56, 1998.

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.
- **Static models of dynamic games**: A game model represents how a player will and would change her beliefs if her opponents take the game in various directions.

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.
- **Static models of dynamic games**: A game model represents how a player will and would change her beliefs if her opponents take the game in various directions.
- A **conditional choice** (do a if E) is rational iff doing a would be rational if the player were to learn E.

Why go beyond the basic Bayesian model?

- **Counterfactual beliefs/choices** are important for assessing the rationality of a strategy.
- **Static models of dynamic games**: A game model represents how a player will and would change her beliefs if her opponents take the game in various directions.
- A **conditional choice** (do *a* if *E*) is rational iff doing *a* would be rational if the player were to learn *E*.
- Strategy choices should be robust against **mistaken beliefs** (weak dominance).

In a game model $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$, different states represent different beliefs only when the agent is doing something different.

$$P_{i,w}(E) = P_i(E \mid [\mathbf{s}_i(w)])$$

To represent different *explanations* (i.e., beliefs) for the same strategy choice, we would need a set of models $\{\mathcal{M}_1^G, \mathcal{M}_2^G, \ldots\}$.

In a game model $\mathcal{M}^G = \langle W, \{P_i\}_{i\in N}, \mathbf{s}\rangle$, different states represent different beliefs only when the agent is doing something different.

$$P_{i,w}(E) = P_i(E \mid B_{i,w}), \quad B_{i,w} \subseteq [\mathbf{s}_i(w)]$$

To represent different *explanations* (i.e., beliefs) for the same strategy choice, we would need a set of models $\{\mathcal{M}^G_1, \mathcal{M}^G_2, \ldots\}$.

In a game model $\mathcal{M}^G = \langle W, \{P_i\}_{i \in N}, \mathbf{s} \rangle$, different states represent different beliefs only when the agent is doing something different.

$$P_{i,w}(E) = P_i(E \mid B_{i,w}), \quad B_{i,w} \subseteq [\mathbf{s}_i(w)]$$

Two way to change beliefs: $P_i(\cdot \mid E \cap B_{i,w})$ and $P_i(\cdot \mid B'_{i,w})$ (conditioning on 0 events).

Game Models

Richer models of a game: lexicographic probabilities, conditional probability systems, non-standard probabilities, plausibility models, . . . (type spaces)

Game Models

Richer models of a game: lexicographic probabilities, conditional probability systems, non-standard probabilities, plausibility models, . . . (type spaces)

"The aim in giving the general definition of a model is not to propose an original explanatory hypothesis, or any explanatory hypothesis, for the behavior of players in games, but only to provide a descriptive framework for the representation of considerations that are relevant to such explanations, a framework that is as *general* and as *neutral* as we can make it." (pg. 35)

R. Stalnaker. *Knowledge, Belief and Counterfactual Reasoning in Games*. Economics and Philosophy, 12(1), pgs. 133 - 163, 1996.

Richer models of games

1. A partition $\approx_i$ representing the different "**types**" of player *i*: $w \approx_i v$ means that *w* and *v* are subjectively indistinguishable to player *i* (*i*'s beliefs, knowledge, and conditional beliefs are the same in both states).

2. A pseudo-partition R_i (serial, transitive and Euclidean relation) representing a player *i*'s **working hypotheses** (full beliefs?, serious possibilities?,...).

3. Player *i*'s **belief revision policy** described in terms of *i*'s conditional beliefs.

Richer models of games

1. A partition $\approx_i$ representing the different "**types**" of player i: $w \approx_i v$ means that w and v are subjectively indistinguishable to player i (i's beliefs, knowledge, and conditional beliefs are the same in both states).

2. A pseudo-partition R_i (serial, transitive and Euclidean relation) representing a player i's **working hypotheses** (full beliefs?, serious possibilities?,...).

3. Player i's **belief revision policy** described in terms of i's conditional beliefs.

This can all be represented by a single relation $\preceq_i \subseteq W \times W$

Richer models of games

$\mathcal{M}^G = \langle W, \{\preceq_i, P_i\}_{i\in N}, \mathbf{s}\rangle$, where W, P_i and **s** are as before and $\preceq_i$ is a reflexive, transitive and locally-connected relation.

1. $w \approx_i v$ iff $w \preceq_i v$ or $v \preceq_i w$. Let $[w]_{\approx_i} = \{v \mid w \approx_i v\}$
2. $w\ R_i\ v$ iff $v \in Max_{\preceq_i}([w]_{\approx_i})$
3. $B_{i,w}(F) = Max_{\preceq_i}(F \cap [w]_{\approx_i})$

   $P_{i,w}(E \mid F) = P_i(E \mid B_{i,w}(F))$

Belief Revision in Games

Exactly how a player revises her beliefs during the game depends, in part, on how she interprets the observed moves of her opponents.

Interpreting Moves

How should Bob respond given evidence that Ann's moves seem irrational?

Interpreting Moves

How should Bob respond given evidence that Ann's moves seem irrational?

1. Bob's beliefs about Ann's perception of the game are incorrect.
2. Bob's assumption about Ann's decision procedure is incorrect.
3. Bob's belief about Ann's assumptions about him is incorrect.
4. Ann's moves are an attempt to influence Bob's behavior in the game.

Interpreting Moves

How should Bob respond given evidence that Ann's moves seem irrational?

1. Bob's beliefs about Ann's perception of the game are incorrect.
2. Bob's assumption about Ann's decision procedure is incorrect.
3. Bob's belief about Ann's assumptions about him is incorrect.
4. Ann's moves are an attempt to influence Bob's behavior in the game.
5. Ann simply failed to successfully implement her adopted strategy, i.e., Ann made a "trembling hand mistake".

Taking Stock

- Game models can be used to *characterize* different solution concepts (e.g., iterated strict dominance, iterated weak dominance, Nash equilibrium, correlated equilibrium, backward induction, extensive-form rationalizability,...)

- Focusing on **dynamic games with simultaneous moves**:

Taking Stock

- Game models can be used to *characterize* different solution concepts (e.g., iterated strict dominance, iterated weak dominance, Nash equilibrium, correlated equilibrium, backward induction, extensive-form rationalizability,...)

- Focusing on **dynamic games with simultaneous moves**:
  - strategy choice is not an instantaneous commitment, but, rather, a representation of what the players will and would do in the course of playing the game
  - Epistemic models describe how the players will and would revise her beliefs during a play of the game

Backward and Forward Induction

There are many epistemic characterizations (Aumann, Stalnaker, Battigalli & Siniscalchi, Friedenberg & Siniscalchi, Perea, Baltag & Smets, Bonanno, van Benthem,...)

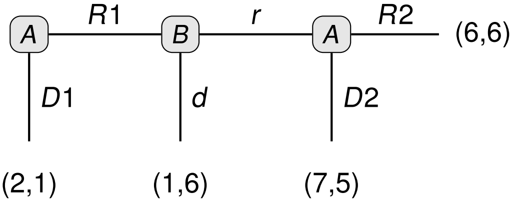
Backward and Forward Induction

There are many epistemic characterizations (Aumann, Stalnaker, Battigalli & Siniscalchi, Friedenberg & Siniscalchi, Perea, Baltag & Smets, Bonanno, van Benthem,...)

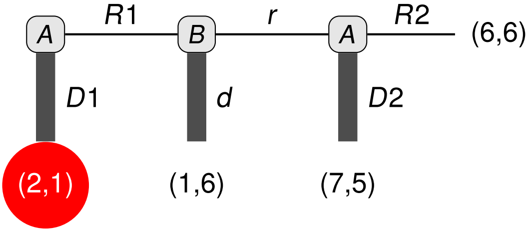
- How should we compare the two “styles of reasoning” about games? (Heifetz & Perea, Reny, Battigalli & Siniscalchi)
- How do (should) players choose between the two different styles of reasoning about games? (Perea, EP & Knoks)

  Aumann & Dreze: *“When all is said and done, how should we play and what should we expect”*.

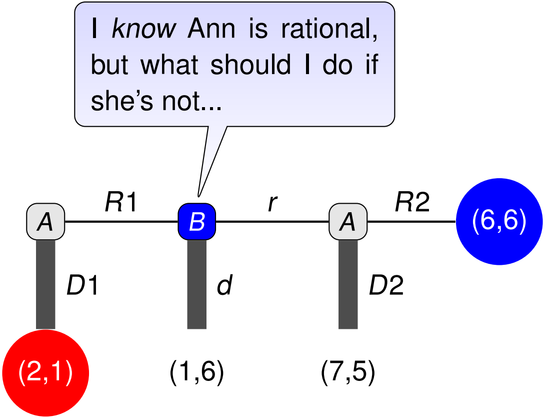
BI Puzzle?

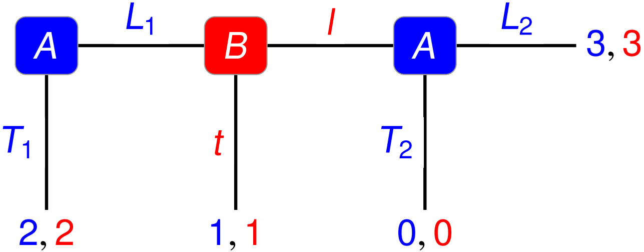


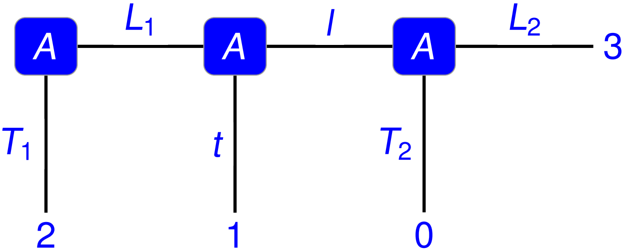
BI Puzzle?

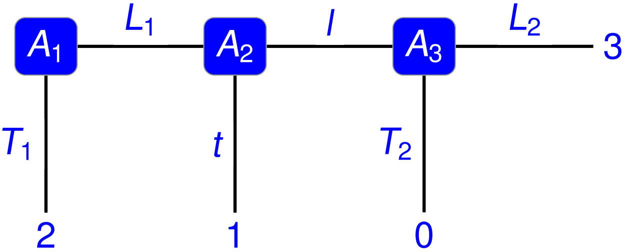


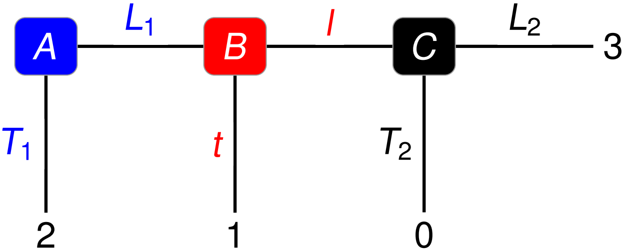
BI Puzzle?

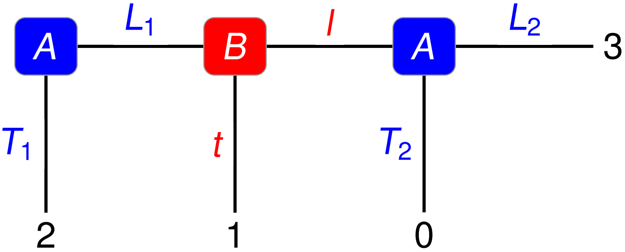

I *know* Ann is rational, but what should I do if she's not...
A
R1
B
r
A
R2
(6,6)
D1
d
D2
(2,1)
(1,6)
(7,5)


A
L_1
B
l
A
L_2
3, 3
T_1
t
T_2
2, 2
1, 1
0, 0

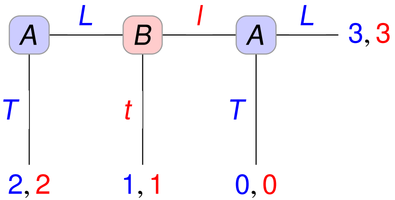

A
A
A
L_1
l
L_2
3
T_1
t
T_2
2
1
0


A_1
L_1
A_2
l
A_3
L_2
3
T_1
t
T_2
2
1
0


A
L₁
B
l
C
L₂
3
T₁
t
T₂
2
1
0


A
L_1
B
l
A
L_2
3
T_1
t
T_2
2
1
0

		Bob	
		t	*l*
Ann	*T*	2,2	2,2
	LT	1,1	0,0
	LL	1,1	3,3



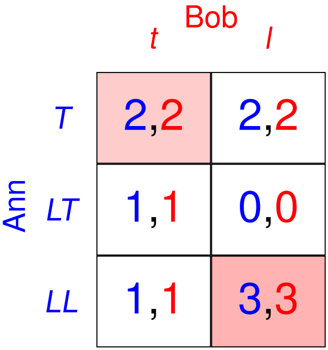
Materially Rational: every choice actually made is optimal (i.e., maximizes subjective expected utility).

Substantively Rational: the player is materially rational and, in addition, for each *possible* choice, the player *would* have chosen rationally if she had had the opportunity to choose.

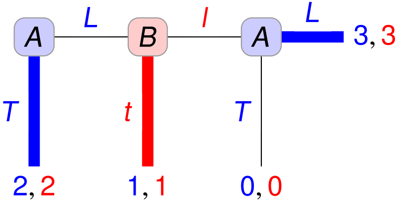
Materially Rational: every choice actually made is optimal (i.e., maximizes subjective expected utility).

Substantively Rational: the player is materially rational and, in addition, for each *possible* choice, the player *would* have chosen rationally if she had had the opportunity to choose.

E.g., Taking keys away from someone who is drunk.



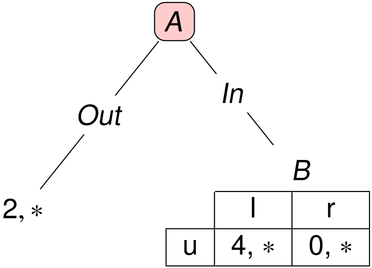
		Bob	
		t	*l*
Ann	*T*	2,2	2,2
	LT	1,1	0,0
	LL	1,1	3,3



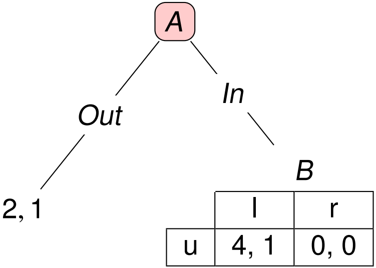
- Bob's belief in a causal counterfactual: Ann would choose *L* on her second move *if* she had a chance to move.
- But we need to ask what would Bob believe about Ann *if* he learned that he was wrong about her first choice. This is a question about Bob's belief revision policy.

Informal characterizations of BI

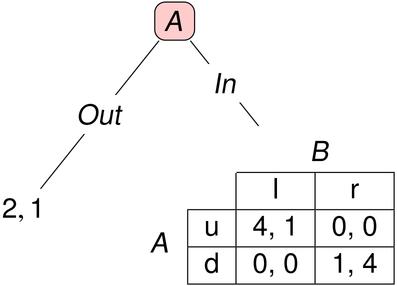
- Future choices are *epistemically independent* of any observed behavior
- Any “off-equilibrium” choice is interpreted simply as a mistake (which will not be repeated)
- At each choice point in a game, the players only reason about future paths



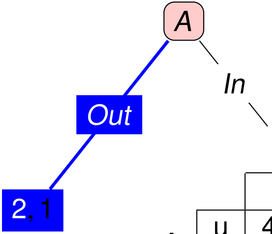
	l	r
u	4, ∗	0, ∗



	l	r
u	4, 1	0, 0



	l	r
u	4, 1	0, 0
d	0, 0	1, 4



		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

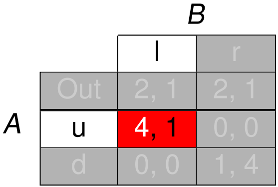
		B	
		l	r
	Out	2, 1	2, 1
A	u	4, 1	0, 0
	d	0, 0	1, 4

		B	
		l	r
	Out	2, 1	2, 1
A	u	4, 1	0, 0
	d	0, 0	1, 4

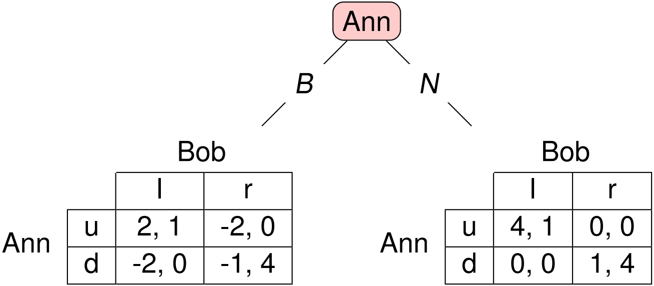
B

		l	r
A	*Out*	2, 1	2, 1
	u	4, 1	0, 0
	d	0, 0	1, 4

		B	
		l	r
	Out	2, 1	2, 1
A	u	4, 1	0, 0
	d	0, 0	1, 4



		B	
		l	r
	Out	2, 1	2, 1
A	u	4, 1	0, 0
	d	0, 0	1, 4



		Bob	
		l	r
Ann	u	2, 1	-2, 0
	d	-2, 0	-1, 4

		Bob	
		l	r
Ann	u	4, 1	0, 0
	d	0, 0	1, 4

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Bob

Ann

	ll	lr	rl	rr
Bu	2, 1	2, 1	-2, 0	-2, 0
Bd	-2, 0	-2, 0	-1, 4	-1, 4
Nu	4, 1	0, 0	4, 1	0, 0
Nd	0, 0	1, 4	0, 0	1, 4

Bob

Ann

	ll	lr	rl	rr
Bu	2, 1	2, 1	-2, 0	-2, 0
Bd	-2, 0	-2, 0	-1, 4	-1, 4
Nu	4, 1	0, 0	4, 1	0, 0
Nd	0, 0	1, 4	0, 0	1, 4

Bob

		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Bob

		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Bob

Ann

	ll	lr	rl	rr
Bu	2, 1	2, 1	-2, 0	-2, 0
Bd	-2, 0	-2, 0	-1, 4	-1, 4
Nu	4, 1	0, 0	4, 1	0, 0
Nd	0, 0	1, 4	0, 0	1, 4

Bob

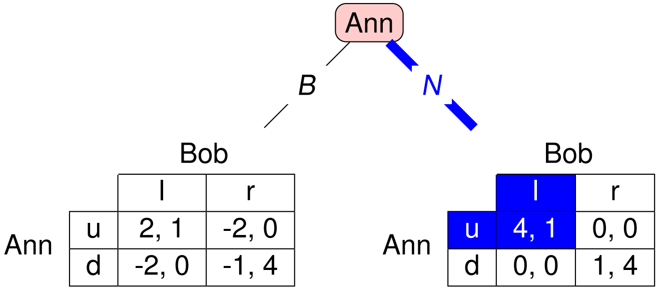
Ann

	ll	lr	rl	rr
Bu	2, 1	2, 1	-2, 0	-2, 0
Bd	-2, 0	-2, 0	-1, 4	-1, 4
Nu	4, 1	0, 0	4, 1	0, 0
Nd	0, 0	1, 4	0, 0	1, 4

Bob

Ann

	ll	lr	rl	rr
Bu	2, 1	2, 1	-2, 0	-2, 0
Bd	-2, 0	-2, 0	-1, 4	-1, 4
Nu	4, 1	0, 0	4, 1	0, 0
Nd	0, 0	1, 4	0, 0	1, 4



Bob

		l	r
Ann	u	2, 1	-2, 0
	d	-2, 0	-1, 4

Bob

		l	r
Ann	u	4, 1	0, 0
	d	0, 0	1, 4

What is forward induction reasoning?

Forward Induction Principle: a player should use all information she acquired about her opponents' past behavior in order to improve her prediction of their future simultaneous and past (unobserved) behavior, relying on the assumption that they are rational.

P. Battigalli. *On Rationalizability in Extensive Games*. Journal of Economic Theory, 74, pgs. 40 - 61, 1997.

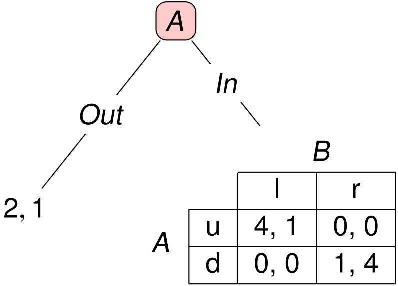
Four key issues

- Should the analysis take place on the tree or the matrix? (plans vs. strategies)
- The players' conditional beliefs must be *rich enough* to employ the forward induction principle.
- Do the players robustly believe the forward induction principle?
- Can players become more/less confident in the forward induction principle?

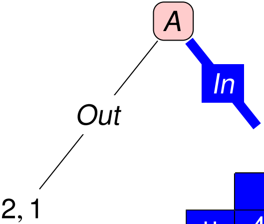
“...in general, a player’s beliefs about what another player will do are based on an inference from two other kinds of beliefs: beliefs about the passive beliefs of that player, and beliefs about her rationality.

“...in general, a player’s beliefs about what another player will do are based on an inference from two other kinds of beliefs: beliefs about the passive beliefs of that player, and beliefs about her rationality. If one’s prediction based on these beliefs is defeated, one must choose whether to revise one’s belief about the other players’s beliefs or one’s belief that she is rational...

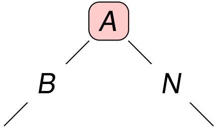
“...in general, a player’s beliefs about what another player will do are based on an inference from two other kinds of beliefs: beliefs about the passive beliefs of that player, and beliefs about her rationality. If one’s prediction based on these beliefs is defeated, one must choose whether to revise one’s belief about the other players’s beliefs or one’s belief that she is rational...But the assumption that the rationalization principle is common belief is itself an assumption about the passive beliefs of other players, and so it is itself something that (according to the principle) might have to be given up in the face of surprising behavioral information. So the rationalization principle undermines its own stability.” (pg. 51, Stalnaker)



		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4



		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

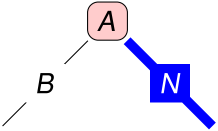


Bob

		l	r
Ann	u	2, 1	-2, 0
	d	-2, 0	-1, 4

Bob

		l	r
Ann	u	4, 1	0, 0
	d	0, 0	1, 4



		Bob	
		l	r
Ann	u	2, 1	-2, 0
	d	-2, 0	-1, 4

		Bob	
		l	r
Ann	u	4, 1	0, 0
	d	0, 0	1, 4

“...Only if one assumes a specific infinite hierarchy of belief revision priorities can one be sure that unlimited iteration of forward induction reasoning will work....But it seems to me that such detailed assumptions about belief revision policy....have no intuitive plausibility.”

(Stalnaker, pg. 53)

Algorithm and a “Theorem”

Algorithm: Eliminate weakly dominated strategies for *just two* rounds, and then eliminate *strictly* dominated strategies iteratively.

Algorithm and a “Theorem”

Algorithm: Eliminate weakly dominated strategies for *just two* rounds, and then eliminate *strictly* dominated strategies iteratively.

“Theorem”: It can be proved that all and only strategies that survive this process are realizable in sufficiently rich models in which it is common belief that all players are rational, and that all revise their beliefs in conformity with the rationalization principle.

Algorithm and a “Theorem”

Algorithm: Eliminate weakly dominated strategies for *just two* rounds, and then eliminate *strictly* dominated strategies iteratively.

“Theorem”: It can be proved that all and only strategies that survive this process are realizable in sufficiently rich models in which it is common belief that all players are rational, and that all revise their beliefs in conformity with the rationalization principle.

Algorithm and a “Theorem”

Algorithm: Eliminate weakly dominated strategies for *just two* rounds, and then eliminate *strictly* dominated strategies iteratively.

“Theorem”: It can be proved that all and only strategies that survive this process are realizable in sufficiently rich models in which it is common belief that all players are rational, and that all revise their beliefs in conformity with the rationalization principle.

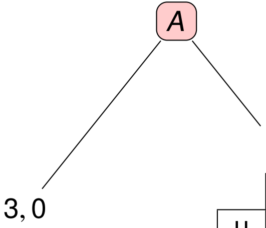
Algorithm and a “Theorem”

Algorithm: Eliminate weakly dominated strategies for *just two* rounds, and then eliminate *strictly* dominated strategies iteratively.

“Theorem”: It can be proved that all and only strategies that survive this process are realizable in sufficiently rich models in which it is common belief that all players are rational, and that all revise their beliefs in conformity with the rationalization principle.

Joint work with Aleks Knoks: “Theorem” $\hookrightarrow$ Theorem

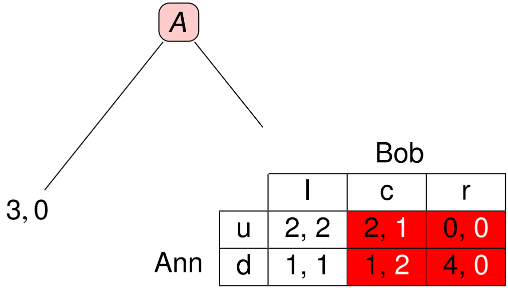
Backward *versus* Forward Induction



		Bob		
		l	c	r
Ann	u	2, 2	2, 1	0, 0
	d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

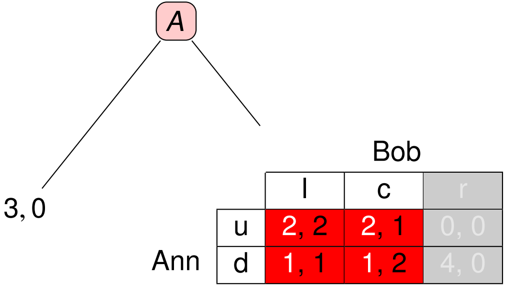
Backward *versus* Forward Induction



	l	c	r
u	2, 2	2, 1	0, 0
d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

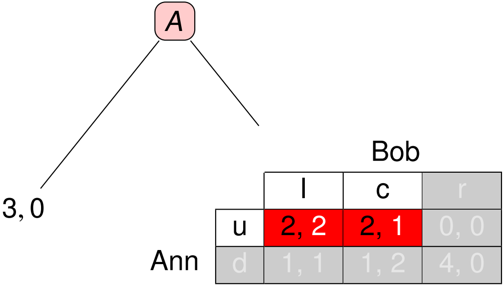
Backward *versus* Forward Induction



			Bob	
		l	c	r
	u	2, 2	2, 1	0, 0
Ann	d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

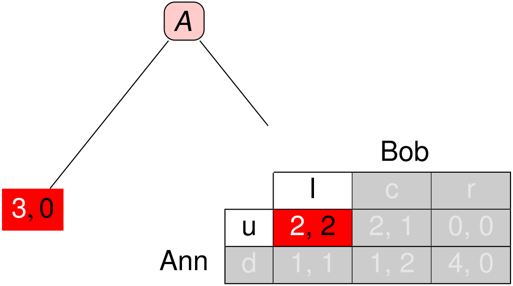
Backward *versus* Forward Induction



	l	c	r
u	2, 2	2, 1	0, 0
d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

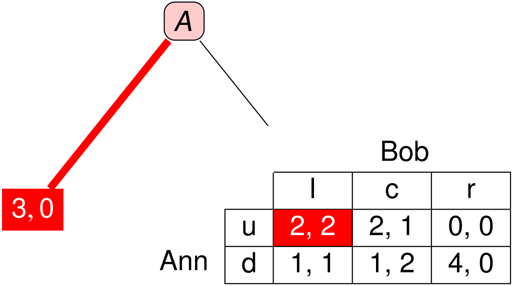
Backward *versus* Forward Induction



	l	c	r
u	2, 2	2, 1	0, 0
d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

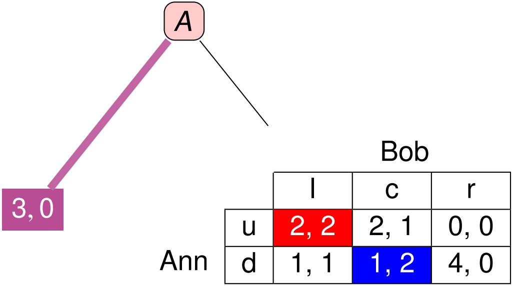
Backward *versus* Forward Induction



		Bob		
		l	c	r
Ann	u	2, 2	2, 1	0, 0
	d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

Backward *versus* Forward Induction

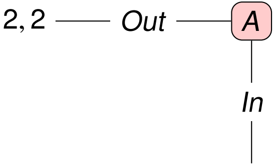


Bob

		l	c	r
	u	2, 2	2, 1	0, 0
Ann	d	1, 1	1, 2	4, 0

A. Perea. *Backward Induction* versus *Forward Induction Reasoning*. Games, 1, pgs. 168 - 188, 2010.

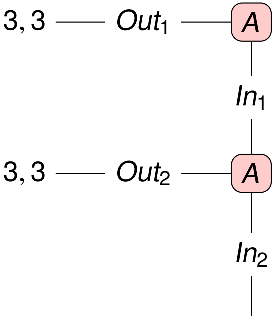
Rationalization *versus* Mistakes



B

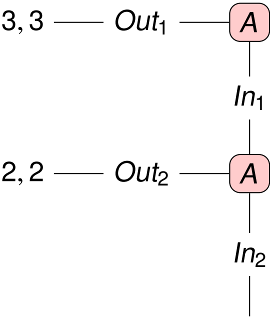
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

Rationalization *versus* Mistakes



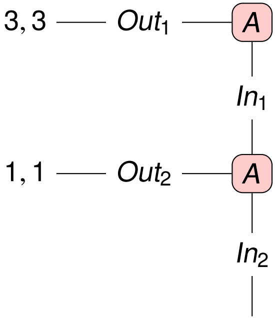
		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

Rationalization *versus* Mistakes



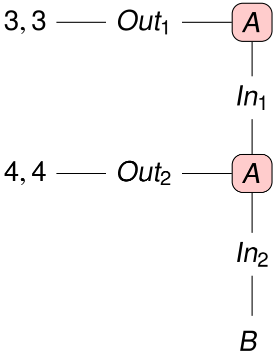
		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

Rationalization *versus* Mistakes



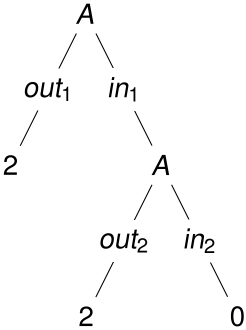
		B	
		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

Rationalization *versus* Mistakes

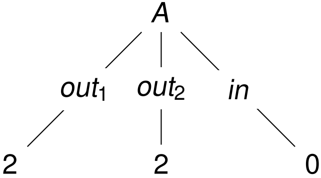


		l	r
A	u	4, 1	0, 0
	d	0, 0	1, 4

Allowing for mistakes



D_1



D_2

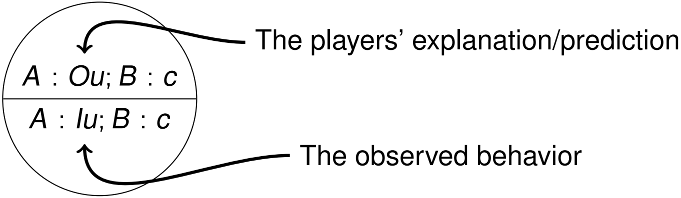
A. Knoks and EP. *Interpreting Mistakes in Games: From Beliefs about Mistakes to Mistaken Beliefs*. manuscript, 2016.

Our Model

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i \in N}, \{\succeq_i\}_{i \in N}, \{P_i\}_{i \in N} \rangle$$

Our Model

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$$



$A : Ou; B : c$

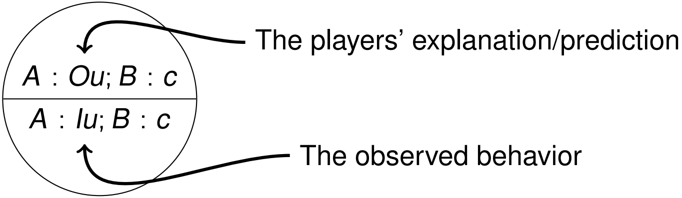
$A : Iu; B : c$

The players' explanation/prediction

The observed behavior

Our Model

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i \in N}, \{\succeq_i\}_{i \in N}, \{P_i\}_{i \in N} \rangle$$



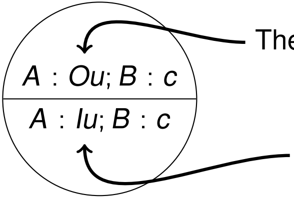
The players' explanation/prediction

The observed behavior

- two functions: β and σ

Our Model

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$$



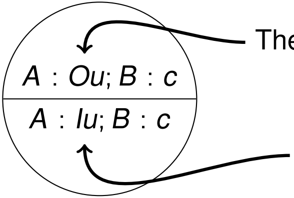
The players' explanation/prediction

The observed behavior

- two functions: β and σ
- states represent *ex interim* stages of the game

Our Model

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$$



The players' explanation/prediction

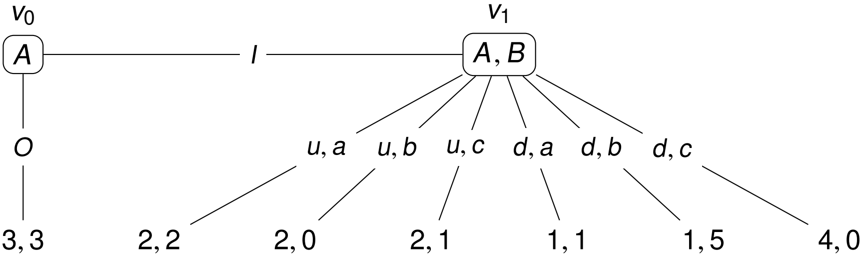
The observed behavior

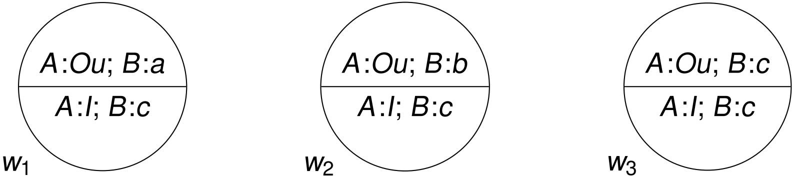
- two functions: β and σ
- states represent *ex interim* stages of the game
- what is a "mistake"?

Suppose that W is a nonempty set of states. Each player i will be associated with two functions β_i and σ_i subject to the following constraints:

1. For each $i \in N$, $\beta_i(w)$ is a (possibly empty) i-history and $\sigma_i(w)$ is a strategy for player i.
2. The i-histories $\{\beta_i(w)\}_{i \in N}$ are **coherent**.

A player made a **mistake at a history** $h \in V_i$ in w provided that $\beta_i(w)_h \neq \sigma_i(w)(h)$ (if $\beta_i(w)_h$ is defined).


v_0
A
I
v_1
A, B
O
u, a
u, b
u, c
d, a
d, b
d, c
3, 3
2, 2
2, 0
2, 1
1, 1
1, 5
4, 0



$A{:}Ou; B{:}a$

$A{:}I; B{:}c$

w_1

$A{:}Ou; B{:}b$

$A{:}I; B{:}c$

w_2

$A{:}Ou; B{:}c$

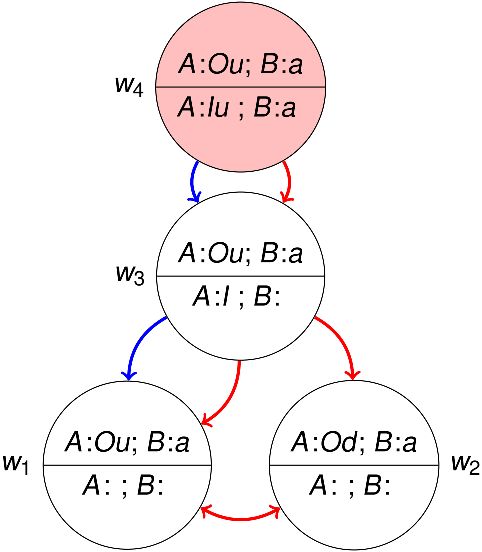
$A{:}I; B{:}c$

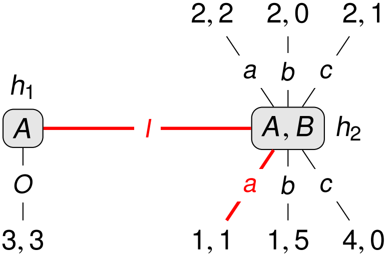
w_3

For a game G, a game model is a tuple $\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i \in N}, \{\succeq_i\}_{i \in N}, \{P_i\}_{i \in N}\rangle$, where:

- For all $w \in W$ and $i \in N$, if $v \in [w]_i$, then $\sigma_i(w) = \sigma_i(v)$. That is, players *know* their own strategy.
- For all $w \in W$ and $i \in N$, for each initial segment $h' \subseteq h_w$ (including the empty history), there is a $w' \in [w]_i$ such that $h_w = h'$.

The Model: Example





Beliefs

$$[w]_i = \{v \mid w \preceq_i v \text{ or } v \preceq_i w\}$$

Beliefs

$$[w]_i = \{v \mid w \preceq_i v \text{ or } v \preceq_i w\}$$

$$P_{i,w}(E) = P_i(E \mid \max_{\succeq_i}([w]_i))$$

Given *any* evidence $F \subseteq W$:

$$P_{i,w}(E \mid F) = P_i(E \mid \max_{\succeq_i}(F \cap [w]_i)).$$

Each state $w \in W$ is associated with a history h_w corresponding to the play of game associated with $\{\beta_i(w)\}_{i \in N}$.

Each state $w \in W$ is associated with a history h_w corresponding to the play of game associated with $\{\beta_i(w)\}_{i \in N}$.

Note that h_w need not be a maximal history, so h_w is the behavior that is observed at state w.

Each state $w \in W$ is associated with a history h_w corresponding to the play of game associated with $\{\beta_i(w)\}_{i \in N}$.

Note that h_w need not be a maximal history, so h_w is the behavior that is observed at state w.

For any $h \in H$, let $[h] = \{w \mid \beta_i(w) = beh_i(h) \text{ for all } i \in N\}$ be the event that the players behaved according to history h.

Each state $w \in W$ is associated with a history h_w corresponding to the play of game associated with $\{\beta_i(w)\}_{i \in N}$.

Note that h_w need not be a maximal history, so h_w is the behavior that is observed at state w.

For any $h \in H$, let $[h] = \{w \mid \beta_i(w) = beh_i(h) \text{ for all } i \in N\}$ be the event that the players behaved according to history h.

$P_{i,w}(E \mid [h_w])$ is i's probability of E given her most plausible explanation of the actions she observed at state w.

Optimal Choice - Induced Strategy

For each $w \in W$, the **strategy realized at** w **by player** i is $s_i(w) : V_i \to Act_i$ defined as follows:

$$s_i(w)(h) = \begin{cases} \beta_i(w)_h & \text{if } \beta_i(w)_h \text{ is defined} \\ \sigma_i(w)(h) & \text{otherwise} \end{cases}$$

Then, $\mathbf{s}(w) = (s_1(w), \ldots, s_n(w))$ is a profile of strategies, and let $Out(\mathbf{s})$ be the (unique) terminal history generated by $\mathbf{s}$.

Optimal Choice - Expected Utility

For any strategy $s_i \in S_i$ for player i, the **expected utility** of s_i at state w is:

$$EU_{i,w}(s_i) = \sum_{w' \in W} P_{i,w}(\{w'\} \mid [h_w])u_i(Out(s_i, \mathbf{s}_{-i}(w))).$$

Optimal Choice

Let $S_i(w) \subseteq S_i$ be the set of strategies for player i that conform to player i's moves in state w.

$$\begin{array}{rcl} Opt_i & = & \{w \mid \sigma_i(w) \text{ maximizes expected utility} \\ & & \text{with respect to } P_{i,w} \text{ and } S_i(w)\}. \end{array}$$

Rational Choice

We say that a state $w' \in [w]_i$ is an **earlier choice state** provided $\beta_i(w')$ is an initial segment of $\beta_i(w)$.

Player i is **rational-1** at state w provided $w' \in Opt_i$ for *all* earlier choice states w'. Let Rat_i^1 be the set of all states w such that i is rational-1 in w.

What can we do with our model?

What can we do with our model?

- Represent FI, BI, as well as “hybrid” belief revision policies

What can we do with our model?

- Represent FI, BI, as well as “hybrid” belief revision policies
- Condition on choices, “mistakes”, observed behavior, rationality, etc.

What can we do with our model?

- Represent FI, BI, as well as “hybrid” belief revision policies
- Condition on choices, “mistakes”, observed behavior, rationality, etc.
- Prove (or re-prove) “characterization results”

What can we do with our model?

- Represent FI, BI, as well as “hybrid” belief revision policies
- Condition on choices, “mistakes”, observed behavior, rationality, etc.
- Prove (or re-prove) “characterization results”
- An implicit assumption in EGT literature: choice of strategy implies its execution: Our framework highlights the role that this assumption plays in (epistemic) game-theoretic analyses

Concluding Remarks: Irrationality in Games

> [T]he rules of rational behavior must provide definitely for the possibility of irrational conduct on the part of others. Imagine that we have discovered a set of rules for all participants—to be termed as “optimal” or “rational”—each of which is indeed optimal provided that the other participants conform. Then the question remains as to what will happen if some of the participants do not conform.
>
> *(von Neumann & Morgenstern, pg. 32)*

Concluding Remarks: Richness Conditions

Many epistemic characterization results make a *richness* assumption about the epistemic models.

- What is a "good" epistemic characterization result?
- ✓ Players need "enough" conditional beliefs to "make sense of" observed behavior.

P. Battigalli and M. Siniscalchi. *Strong Belief and Forward Induction Reasoning*. Journal of Economic Theory 106(2), pgs. 356391, 2002.

R. Stalnaker. *Belief revision in games: forward and backward induction*. Mathematical Social Sciences, 36, pgs. 31 - 56, 1998.

Concluding Remarks: Reasoning in Games

"The word *eductive* will be used to describe a dynamic process by means of which equilibrium is achieved through careful reasoning on the part of the players. Such reasoning will usually require an attempt to simulate the reasoning processes of the other players. Some measure of pre-play communication is therefore implied, although this need not be explicit. To reason along the lines "if I think that he thinks that I think..." requires that information be available on how an opponent thinks."

(pg. 184)

K. Binmore. *Modeling Rational Players*. Economics and Philosophy, 3,179 - 21, 1987.

Concluding Remarks: Reasoning in Games

"*The* fundamental insight of game theory [is] that a rational player must take into account that the players reason about each other in deciding how to play" (pg. 81)

R. Aumann and J. Dreze. *Rational expectations in games*. American Economic Review, Vol. 98, pgs. 72 – 86 (2008).

Concluding Remarks: Reasoning in Games

"*The* fundamental insight of game theory [is] that a rational player must take into account that the players reason about each other in deciding how to play" (pg. 81)

R. Aumann and J. Dreze. *Rational expectations in games*. American Economic Review, Vol. 98, pgs. 72 – 86 (2008).

Exactly *how* the players incorporate the fact that they are interacting with other (actively reasoning) rational agents is the subject of much debate.

Thank You!